DOCUMENT RESUME

ED 413 344 TM 027 640

AUTHOR Sinkule, Judith Louise
TITLE A Study of Standardized Tests and Alternatives for Elementary Schools.
PUB DATE 1996-12-00
NOTE 79p.; Master's thesis, Salem-Teikyo University.
PUB TYPE Dissertations/Theses - Masters Theses (042)
EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Elementary Education; *Elementary School Teachers; Interviews; Principals; *Standardized Tests; *Student Evaluation; Surveys; *Teacher Attitudes; *Test Use
IDENTIFIERS *Alternative Assessment; Randolph County School District WV

ABSTRACT

A review of the literature on the use of standardized tests suggested that standardized testing has not improved education and that alternative assessments are a more accurate measure of a student's knowledge. A study conducted in Randolph County (West Virginia) investigated whether elementary school teachers agreed with these conclusions. Results from 71 elementary school teacher surveys and interviews with a county office administrator and 2 principals indicated that all participants saw a need for standardized tests, even though most disliked using them. Teachers did use alternative assessments, but considered standardized testing to be of equal importance. In fact, teachers thought that the use of standardized test results, in conjunction with alternative assessment results, would most accurately measure students' educational progress. Overall, teachers in Randolph County thought that standardized tests should not be used alone in student evaluation. Nine appendixes provide supplemental information, including the examples of test questions for students, the teacher survey, and the interview protocols. (Contains 13 figures and 47 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

TM

ED 413 344

A STUDY OF STANDARDIZED TESTS AND ALTERNATIVES FOR ELEMENTARY SCHOOLS

A Thesis

presented to

the Faculty of the Graduate School

Salem-Teikyo University

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Judith Sinkule

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

In Partial Fulfillment

of the Requirements for the Degree

Master of Education

TM027640

by

Judith Louise Sinkule

BEST COPY AVAILABLE

December 1996

Salem-Teikyo University

Salem, West Virginia

This thesis submitted by Judith Louise Sinkule has been approved as meeting the research requirements for the Master of Education degree.

1/14/97 _____ **Research Advisor**
Administrator for Graduate Education and Assistant Professor of Education, Salem-Teikyo University, Salem, West Virginia

1/14/97 _____ **Graduate Study Committee Member**
Assistant Professor of Education, Salem-Teikyo University, Salem, West Virginia

1/14/97 _____ **Graduate Study Committee Member**
Professor of Education, Salem-Teikyo University, Salem, West Virginia

c 1996

Judith Louise Sinkule

ALL RIGHTS RESERVED

ACKNOWLEDGMENTS

I wish to thank several people whom without their help and perseverance, I would never have completed this thesis. First of all, I wish to thank Dr. Sharon Brisbin for her affable supportiveness and confidence in my ability throughout the completion of my work. Her willingness to help at any time was of great assistance to me.

My sincere appreciation is extended to Principal Ferdinand Baldt and Principal Rebecca Whiteman, who provided helpful information on the subject of standardized testing from their viewpoint. I must also offer my gratitude to Mr. Vic Thompson, Director of Student Services for Randolph County. He provided me with administrative reasoning and state mandates for testing.

I thank the elementary school teachers of Randolph County who took time out of their busy schedules to answer my survey. Without their help, I could not have compiled statistics for my thesis. Their perspectives were invaluable tools without which I could not have reached any conclusion

Last, but not least, I thank my family--my husband, Ron and my children, Jenny and John. They persevered while I worked on this project, waiting their turn on the computer while I spent endless hours on it. I could not have completed my Master Degree thesis without their patience and understanding.

ABSTRACT

Should standardized tests be replaced with or used in conjunction with alternative assessments in order to be a more effective measurement tool for evaluating students' abilities in elementary schools? Accountability is a major issue in the subject of standardized testing, and constitutes the importance of this study. Research done by the author of this thesis has indicated that standardized testing has not improved education; and that alternative assessments are a more accurate measure of a student's knowledge. There is a move by some states to alter the format of standardized tests, so that they measure higher order thinking skills, rather than just the memorization of basic facts. The author conducted a study in Randolph County, West Virginia to determine whether or not elementary school teachers in Randolph County agreed with the research. Results from seventy-one elementary school teacher surveys, as well as two principal interviews and one County Office administrator interview, were utilized. Results of the research indicated a need for standardized tests by all participants, even though the majority of participants disliked using standardized tests. Results also revealed that teachers use alternative assessments, but consider standardized test results to be of equal importance. The teachers indicated the use of standardized test results in conjunction with alternative assessment results, to more accurately measure their students' educational progress. Thus, the conclusion of this study of standardized tests versus alternative assessments did indicate that the majority of the participants, including interviewed principals and administrator felt that standardized test results should not be used alone.

TABLE OF CONTENTS

CHAPTER		Page
1.	The Problem: Introduction	1
	Statement of the Problem	2
	Research Questions	3
	Assumptions of the Study	3
	Limitations of the Study	4
	Importance of the Problem	4
	Definition of Terms	4
2.	Related Literature: Introduction	8
	History of Standardized Testing	8
	Types of Testing	10
	Advantages and Disadvantages of Standardized Testing	12
	Alternatives to Standardized Testing	18
	Problems with Alternative Assessments	21
	Current Developments in the Use of Alternative Assessments	23
	Summary	24
3.	Methods and Procedures: Introduction	25
	Collection of Data	26
	Treatment of Data	27

Summary. 28

4. Analysis of Collected Data: Introduction. 29

Presentation of Data. 29

Explanation and Analysis of Data. 42

Summary. 50

5. Summary and Conclusions: Introduction. 51

Conclusions. 51

Recommendations. 52

Summary. 53

APPENDICES

A. Standardized Achievement Test Example. 54

B. Performance Assessment Example. 55

C. Science Alternative Assessment Example 57

D. Writing Alternative Assessment Example. 58

E. Request for Survey to Superintendent. 59

F. Elementary Teacher Survey. 60

G. Permission from Superintendent. 63

H. Principal Interview Questions. 64

I. Director Interview Questions. 66

SELECTED BIBLIOGRAPHY. 67

CHAPTER 1

THE PROBLEM

Introduction

This study has proposed that alternative assessment methods to standardized testing exist, and that they are a more accurate measure of student achievement. However, this study has indicated that there are problems with both types of assessment. One of the major problems of both types discovered was time. The following poem was chosen to illustrate the relevance of the time factor involved in the use of standardized tests. The poem also relates that standardized tests are an instructional interruption.

> This is a Test
> We interrupt our regularly scheduled
> Educational Program for a test
> Of the emergency testing system:
> TAAS, PSAT, TEAMS, MAT NAEP.
>
> If this had been a real test,
> SAT, ACT, TASP, GED,
> You would have been directed
> To a designated testing site.
>
> Repeat: This is only a test
> TASS, PSAT, TEAMS, MAT, NAEP.
> We will now return you to our
> Regular Educational Program (9: 59).

Testing should be used in conjunction with instruction and learning, rather than an interruption of it, as the above poem suggests (22: 29). Otherwise, learning is not adequately effective for the teacher or the student.

This study was conducted within the Randolph County school system. Randolph County has a geographical area of 1,046 square miles, one of the

largest counties east of the Mississippi. This makes it a unique area because of the location of its schools, some rural and others urban (45: np). The county has eleven elementary schools in the county, seven of which are rural. Two of these rural schools include kindergarten through twelfth grades. The twelve year schools are Harman and Pickens. Pickens School is located approximately thirty miles southwest of Elkins on top of a mountain. The remaining eight schools house grades kindergarten through fifth. Beverly Elementary School, Coalton Elementary School, George Ward Elementary School, Homestead Elementary School, and Valley Head Elementary School are rural elementary schools. Valley Head is located approximately thirty-five miles south of Elkins. The four urban elementary schools are Jennings Randolph Elementary School, Midland Elementary School, Third Ward Elementary School, and North Elementary School. The elementary student population of these schools is 2,422 and 108 elementary teachers staff these schools.

All of these elementary schools are at maximum enrollment except for the rural schools of Homestead, Pickens, Harman, and Coalton. Three of the rural schools, George Ward, Homestead, and Valley Head comprise twenty-five percent of the county student population. Well over fifty percent of this student population qualify for free and reduced lunch (45: np).

Statement of the Problem

The Problem. Should standardized tests be replaced with or used in conjunction with alternative assessments in order to be a more effective

measurement tool for evaluating students' abilities in elementary schools?

Research Questions

1. How much time is spent in the classroom on preparation for standardized tests?

2. How much time is spent in the classroom on administration of standarized tests?

3. Do elementary teachers feel compelled to "teach to the test"?

4. How many alternative assessment methods of student evaluation do elementary teachers in Randolph County currently utilize?

5. Can these alternative assessment methods be more effective tools of student evaluation than standardized tests?

Assumptions of the Study

The First Assumption. The measurement tool used for this study was valid.

The Second Assumption. The sample size used for this study was adequate.

The Third Assumption. The sample used in this study was typical of elementary teachers in Randolph County.

The Fourth Assumption. The participating teachers understood the questions on the questionnaire.

Limitations of the Study

The First Limitation. The survey was limited to elementary teachers in Randolph County.

The Second Limitation. The fact that standardized tests must be given in Randolph County cannot be changed.

Importance of the Problem

This study was important because of accountability issues. Teachers are accountable for teaching their students what they need to know in order to succeed in their adult lives. The accountability problem has led to standardized testing of the masses. Teachers are being held responsible for results from standardized tests, despite the fact that standardized tests may not evaluate true abilities of their students. This study was important in discovering new methods of evaluation that will better prepare students for the real world.

Definition of Terms

Accountability. Educationally speaking, the demand that schools account for money spent on education by providing evidence of student learning and achievement (3: 105).

Achievement test. A test designed to measure students' "school taught" learning in such areas as reading, writing, mathematics, and science (3: 105).

Alternative assessment. A range of assessment methods designed to take the place of or to supplement standardized tests (3: 105).

America 2000. This is an education reform plan of the federal government introduced by President George Bush and Secretary of Education Lamar Alexander in April 1991, which contains the four concepts of choice, higher standards, radical reform, and national testing (12: 186).

Assessment. The process of gathering information about students and what the students know and can do (3: 1).

Authentic assessment. This is an assessment that both mirrors and measures students' performance in "real-life" tasks and situations. For example, if the teachers want students to communicate effectively in writing, the authentic way to assess them is to evaluate actual samples of their writing (3: 106).

Competencies test. An assessment designed to make sure that students have met minimal content and skill standards (3: 106).

Content validation. The processes by which an expert, or experts, examine a particular test and provide the equivalent of an imprimatur--a license to print or publish (4: 21).

Criterion-referenced test. An assessment designed to reveal what a student knows, understands, or can do in relation to specific objectives (3: 106).

Evaluation. This is the process of interpreting and making judgments about assessment information (3: 1).

Grade equivalent score. A test score which expresses both the grade and the month of the school year, typically divided into tenths (9: 314).

High stakes. A term used in reference to consequences attached to standardized test results. When the consequences involve monetary or

teacher/student rewards, the test results are said to have high stakes.

Holistic assessment. *See* Authentic assessment (3: 109).

National curriculum. This is an educational curriculum in every school district in the nation that follows the same guidelines and objectives.

Norm. The midpoint in the distribution of scores obtained in testing a large group of students. By definition, fifty percent of tested students score above the norm, and fifty percent score below (3: 110).

Norm-referenced test. A standardized assessment designed to place a student or group of students in rank order compared to other test takers of the same age and grade (3: 110).

Outcome-based assessment. An assessment that is designed to "test to the objectives." Instead of emphasizing what a student has not mastered, outcome-based assessment generally emphasizes what a student has learned (3: 110).

Pedagogy. This is the integration of assessment and instruction (4: 72).

Performance-based assessment. This is a direct, systematic observational assessment based on student performances or performance samples and established performance criteria (3: 111).

Portfolio. A file, folder, computer disk, or box containing information and work samples that document a student's growth and accomplishments over time (3: 111).

Reading comprehension. A thought communication process which involves two primary components--the rate at which the thoughts are received

and the accuracy with which the thoughts are understood (10: 348).

Reading efficiency. This is general reading ability (10: 347).

Reading rate. This is speed (10: 347).

Standardized test. This is a mass-produced multiple-choice test that could be administered to large numbers of people with consistent results (3: 3).

Title I. A federally funded program provided to remediate a subject, usually reading or math.

CHAPTER 2

RELATED LITERATURE

Introduction

This related literature was reviewed to learn about assessment practices, both old and new, at the elementary school level. Twenty-nine of thirty-five articles reviewed in this literature favored new assessment measures for evaluating student achievement. Although new assessment methods are favorable, obstacles still exist. The research in this study will explore new assessment measures and the problems they present.

History of Standardized Testing

The author of a book titled Authentic Assessment, Hart, stated that "testing is an ancient tradition." She conveyed that the Old Testament described a test devised by Jephthah of Gilead to distinguish his own men from enemy troops. The test was a pronunciation of a password, and those who failed the test were executed (4: 2).

Additionally, Rothman, a nationally known writer on educational policy, related that eighteen hundred years later, Horace Mann, a leading educator, wished to reform the ancient tradition of testing. Mann faced opposition to his proposals for school reform and in an effort to make his case stronger, created a test to be given to Boston city school children. Mann reasoned that this novel testing instrument would provide objective information on the current quality of

teaching and learning in Boston schools. The results of this test exposed wide gaps in Boston's students' knowledge, which bolstered Mann's arguments for change (6: 33).

By the late nineteenth century, Rothman reported an interest among policy makers in tests as management tools had grown rapidly. Schools had to justify their performance quantitatively to tax payers, and standardized tests were believed to help them do this more effectively. With their multiple-choice and short answer questions, these tests were cost efficient, which made them attractive tools of evaluation (6: 36).

Following World War I, during the early years of the twentieth century, Rothman reported on the Army Alpha test. It was administered to more than 1.7 million soldiers (6: 37). Stephen Jay Gould, a biologist who wrote a critique of these tests, stated that a technology had been developed for testing all pupils. Tests could now rank and stream everybody. The era of mass testing had begun (3: 195). Hart also added that by the end of World War I, more than one hundred test bureaus were producing standardized tests, and by 1939, twenty-six states had initiated statewide testing programs (4: 3).

Also, according to Hart, two trends contributed to further the test industry after World War II. The first of these was the effect industrialization had on education (4: 3). The idea that schools could run like factories of the 1950's appealed to Americans. Standardized tests complemented the factory-type school, where learning was broken down into small skills and bits of knowledge.

The second trend that helped the test industry grow was the concern about accountability. Student achievement was not keeping pace with the large amount of money being spent on education. Scholastic Aptitude Test scores were falling. Thus, to measure and monitor education more closely, the amount of money spent on testing quadrupled in the United States between 1955 and 1986 (4: 4). During the 1980's, United States children became the most tested students in the world, but the desired improvements in education reform did not occur, announced Monty Neill at a government hearing in 1991 (33: 49). This accountability dilemma led the government, under the Bush Administration, to propose a national exam. The testimony of Monty Neill during the Hearing on the Office of Educational Research and Improvement, stated the desire of child advocates and educational groups is to support efforts to introduce new assessments as part of implementing school reform and accountability, rather than implementing a national test (33:49). Neill stated that school reform is necessary, not more testing. The need is for genuine accountability, not test scores from multiple-choice or short-answer exams, argued these groups in this government report (33: 50).

Types of Testing

Hart described two methods of testing, authentic tests and standardized tests. She defined authentic tests as any activity that involves students in tasks that are worthwhile, significant, and meaningful (4: 9). According to an ERIC

Clearinghouse document published for parents in 1993, standardized tests usually are created by commercial test publishers and usually are designed to give a common measure of students' performance (34: 1). Included in Appendix A is an example of a standardized test.

Hart also contended that most standardized tests fall into two main categories, norm-referenced tests and criterion-referenced tests. Both these tests rely on multiple-choice questions. An example of norm-referenced test scores would be a report comparing student achievement in local schools against a statewide or nationwide norm. Competency tests and achievement tests are examples of criterion-referenced tests. Hart said that generally, students are required to pass a competency test for some type of promotion or graduation (4: 5). She explained that criterion-referenced tests are used to identify strengths and weaknesses in individual students (4: 106).

Likewise, Rothman related that authentic assessment generally falls into three categories, performance-based assessments, projects, and portfolios (6: 71). Examples of authentic assessments are included in Appendices B, C, and D. Lewis explained that performance-based assessment measures how well students can do something with their knowledge. This alternative assessment gages problem-solving skills, rather than simply having students answer questions (23: 73). Denman recounted the teaching of her high school newspaper class as an example of a project assessment. "You cannot get much more authentic than the production of a newspaper . . . " , said Denman (16: 55).

Hart reported a portfolio to be a container of samples of an individual's skills, ideas, interests, and accomplishments (4: 23). In his research concerning performance assessments, Khattri indicated that these three categories of authentic assessment may overlap during the learning process, as portfolios may be used to complete project-based assignments (21: 83).

Advantages and Disadvantages of Standardized Testing

An ERIC Clearinghouse pamphlet written for parents regarding standardized tests stated that standardized tests are designed to give a common measure of students' performance. In this brochure, Bagin and Rudner relayed that standardized tests are tools used by educators to give them a common yardstick or standard of measure. This standard of measure indicates the success or failure of a school program and gives an indication of the skills and abilities of students. In addition, standardized test scores can help teachers and administrators make decisions regarding the current instructional program (34: np).

However, in an article written by an associate professor of educational research and measurement at the University of Toledo in Ohio, Cizek suggested that standardized test scores are often only casually perused and then filed, without any analysis of how the results might be used to improve general or individual instruction (13: 248). Similarly, Campbell, a director of a testing and evaluation service at Oklahoma State University, contended that standardized

tests are important tools in decision making only if scores are correctly interpreted. For example, serious deficiencies exist with grade equivalent scores. Children do not learn in a uniform manner within and between grade levels. This lack of uniformity has an effect on grade equivalent scores (11: 314). Test publishers may use different norm groups or write tests that have more items covering a particular skill than other publishers. These differences in content emphasis cause differences in students' grade equivalent scores (11: 315).

Likewise, Popham agreed that when used properly, test scores can determine whether instruction is meeting its goals (41: 629). He further explained, in this ERIC article, that four competency testing programs had been used in Texas, Detroit, South Carolina, and Maryland. These school systems utilized their test results to improve instruction in the needed subject areas (41: 634).

Moreover, Ebel, a specialist in educational measurement and research, contended that teachers need to use standardized test scores as evidence of their effective teaching (1: 8). Standardized tests provide a basis for justifying their continued employment. He further stated that these types of tests will motivate students and give them direction in their learning (1: 18). Likewise, he remarked, standardized tests will motivate teachers to teach well (1: 31).

In contrast, Hynes, a freelance writer and substitute teacher, pointed out that teachers are teaching their students how to be successful on reading comprehension test questions without reading the accompanying paragraph. He

maintained that students can be taught to choose correct answers to math test questions without doing the calculations (19: 559). He further stated victory in test-taking is a matter of frequent practice in taking similar tests (19: 562).

Furthermore, schools with high enrollments of minority students are more likely to pressure their teachers to raise scores on standardized tests; Bracey, a research psychologist, related in his research concerning the impact of testing on teaching. Teachers under these circumstances, he related, feel pressure to increase test scores. Administrators require these teachers to spend considerable time on preparing students for testing to ensure increased test scores (8: 806).

In the same vein, Darling-Hammond added that using rewards and sanctions for test scores have disadvantages of inequality. Test results may punish minority students by causing more experienced teachers to seek employment in schools where students are easy to teach. She further maintained that neither employability nor earnings are affected by students' scores on standardized tests (15: 223).

Additionally, a study was done by Shepard and Bliem utilizing parent surveys to learn how parents evaluated the usefulness of standardized tests compared to performance assessment. The study revealed that one of the reasons for parents' interest in standardized tests was in finding the normative or comparative information about the quality of education at their child's school (30: 26). However, parents in this study found that talking with their child's teacher

and seeing samples of their child's work was also a reliable means of evaluating the school (30: 27). This study by Shepard and Bliem, indicated that one subgroup of parents preferred standardized tests for mathematics, because they felt that in math only one right answer exists (30: 30).

Furthermore, even parents in Shepard and Bliem's study who preferred standardized tests, often noted that performance assessment was a better tool for higher-order thinking skills (30: 28). In her article concerning accountability, Lieberman reinforced this theory. She stated that standardized tests do not foster the teaching of real-world tasks (24: 221). Standardized tests stifle creativity, and less time is spent teaching thinking and reasoning skills, she said (24: 219).

Meanwhile, researchers and theorists who contribute reading ability to efficiency and rate of reading, as well as comprehension, have questioned the reliability of standardized tests to measure these factors efficiently. Since standardized tests of reading ability are usually timed, indicated Carver, there is the question as to what extent standardized tests of reading ability are actually measuring reading efficiency; or to what extent they measure accuracy and rate, or comprehension (12: 347). Carver recounted his research to determine which of four standardized tests actually measured reading efficiency. The first test was the Iowa Test of Basic Skills--ITBS. Also examined was the Rauding Efficiency Level Test--RELT. Carver defined rauding as the means to accurately comprehend the complete thoughts in sentences as they are being read. The

other two tests used were the Degrees of Reading Power test--DRP, and the Nelson-Denny Reading Test--NDRT (12: 348). Results of Carver's study showed that standardized tests of reading comprehension--such as the ITBS, DRP, and NDRT--measured general reading ability only. The RELT, which should have measured reading efficiency, actually measured reading rate (12: 357).

Subsequently, students experiencing difficulty with reading, will tend to perform poorly on standardized tests in the content areas or portions of the math which require reading, remarked Perrone, a published educator. In this case, a student's true knowledge may be underestimated (5: 26). Similarly, a good reader, he said, may choose a wrong answer on a standardized test because of ambiguities in the selection of multiple-choice answers given (5: 27). Regardless, argued Ebel in asserting his agreement with standardized tests, ambiguity can be minimized with skill and care in test construction (1: 11).

Paralleling this idea of ambiguities, Hynes defined distractors as a set of wrong answers for a multiple-choice question. Changing the distractors, Hynes remarked in his article about test taking, changes the outcome of a test. Hynes maintained that even changing the order in which distractors and the correct answers are arranged, will have an effect on a student's answer (19: 560).

Perrone cited many variables, both behavioral and mechanical, that affect the reliability of standardized tests including the physical and mental health of a child during testing, the noise in a classroom, the attitude of the teacher administering the test, and similar testing experiences of the child (5: 25). A

broken pencil during testing, marking an incorrect box or bubble, overlooking a question, missing a word while reading are all mechanical factors which may account for reliability errors (5: 26). Bagin and Rudner confirmed Perrone's notion inasmuch as a child's scores will vary from day to day, dependent on whether he guesses, receives clear directions, follows directions carefully, takes the test seriously, and is comfortable taking the test (34: np).

Validity has been another concern in the use of standardized tests, according to Perrone in his book about the abuses of standardized tests. He related that content validation, the process by which experts examine a particular test and provide a license to publish, is the focus of most standardized tests used in the elementary school. He further stated that most standardized tests have been constructed to conform to predetermined educational objectives through which students are expected to work. These types of tests have less relationship to programs that promote individualism and flexible educational objectives (5: 21).

In conclusion, Kaufhold, a professor of education at a Florida university, enumerated several reasons not to rely only on standardized tests. He alleged that, in the absence of a national curriculum, school districts emphasize different subject areas. He also pointed out that more economically stable school districts can afford special tutoring programs, more technology, and more teachers, thus increasing standardized test scores. Students also differ greatly in terms of values, expectations, and motivation, making the reliability and validity of test score results more questionable (20: 41). Regarding textbooks as they relate to

subject matter for the same grade level, he found that they vary from school district to school district. Thus, comparison of one district's test scores with another would not be valid if the school districts are emphasizing different subject areas in their curriculum. Additionally, Kaufhold claimed the environment of the more elite schools would be more conducive to higher test scores, simply because of finer facilities. School climate is another variable, he declared. In schools with little discipline, even students who want to learn find it difficult to do so. "Until some order is brought to troubled schools, comparisons for standardized test scores won't be valid," said Kaufhold (20: 42).

Numerous advantages and disadvantages are related to both standardized tests and alternative assessments. Kaufhold made five suggestions to school board members for assessing students: he suggested using locally-constructed standardized tests; taking a baseline measurement of student achievement for comparison to test results; choosing school districts with similarities for comparison; comparing students not grades; and utilizing subjective judgments as well as test results (20: 42).

Alternatives to Standardized Testing

The search for alternatives to standardized tests as a means of evaluation has led to the discovery of authentic assessment, offered Hart. Authentic assessment has many names, including, alternative assessment, performance assessment, holistic assessment, or outcome-based assessment. They are all

new approaches to assessment, but have slightly different emphases, thus the varied nomenclatures. They all imply a movement toward assessment that supports rather than corrupts exemplary teaching (4: 9).

Lewis reported that one of the factors that is urging states to search for assessment alternatives is the revision of the Title I legislation. This legislation states that students who are served by Title I must meet the same standards as other students in the state. Chief among the alternative assessments chosen was performance assessment. This type of assessment might include hands-on science experiments, exhibits, or portfolios (23: 73).

Furthermore, in his article describing assessment practices being used in a Fairfax, Virginia elementary school, Afflerbach reported that assessments proved to be most productive when students were assessed as an integral part of the learning process. Assessments should be chosen by teachers, he said, based on what is meaningful for their students at a particular time (7: 624). He continued that teachers in this Fairfax, Virginia elementary school chose alternative assessments as a means to meet the literacy needs of students. These students ranged from fluent readers and writers to those who had limited or no experience with English. Some of the alternative assessments used in the above study were running records--an observational task whereby teachers note a student's reading errors (7: 622). Another observational task used was anecdotal notes, in which teachers keep a dated written record of any type of observation they deem important to a particular student. Writing vocabulary and

word dictation was another method used by teachers in this study. For this assessment, students were asked to write any words they think they can write and then read these back to the teacher. Interviews with the students gave teachers an idea of the students' feelings about what they were learning and what more they wished to learn about a particular subject. Journals were a final alternative assessment mentioned in this study, which were used by teachers to help them direct their instruction to the students' needs and interests (7: 623).

In agreement with the concept of integration of assessment and instruction is Fred Newmann, Director of The Center on Organization and Restructuring of Schools. Newmann indicated that, for some schools, alternative assessment was a product of restructuring their school. He used the terms authentic instruction and authentic student performance, rather than alternative assessment, in his research (10: 70). In an interview by Brandt for *Educational Leadership* magazine, Newmann defined the integration of assessment and instruction as pedagogy. He explained that authentic student performance has three characteristics. The first characteristic was construction of knowledge, or the method used by the student to successfully analyze and interpret information as opposed to simply reproducing it. The second characteristic was disciplined inquiry, the method used by the student to work within a knowledge base, investigate something, and use communication skills to convey what is learned. The last characteristic of authentic student performance concerned the final product and its value or meaning to the student or to others beyond certifying

mastery in school (10: 72).

Kim also concurred with the issue of integrating teaching, testing, and learning. His paper presented at the annual meeting of the Mid South Educational Research Association in 1992 stated that the effective use of tests for instructional guidance requires the integration of these three educational processes (44: 29). He further stated a need for investigation of the relationships among the major components of teaching, testing, and learning (22: 30).

Furthermore, Maeroff asserted alternative assessment does not drive instruction, as do standardized tests, but grows out of and is part of the curriculum (38: 274), yet there are also drawbacks to this method of assessment, reported Maeroff. Alternative assessments are time-consuming, labor-intensive, and imprecise, and there is no means to assess large numbers of students at one time (38: 275).

Meanwhile, according to Shepard in her article for *Educational Leadership*, introducing performance assessments aimed at thinking and problem-solving should improve instruction and learning (35: 1). Her study, done with third grade teachers, proved that when faced with new challenges, effective teachers found ways to overcome the challenges, thus improving their teaching and their students' learning (35: 4).

Problems with Alternative Assessments

Alternative assessments may be a more desirable type of assessment for

some, yet there are still disadvantages. The most significant disadvantages involve time and money (38: 275). Many articles confirmed this hypothesis.

In his article regarding alternative assessment, Maeroff reported some conclusions drawn by measurement experts from the United States after their meeting in March 1991 with the federal government's Center for Research on Evaluation, Standards and Student Testing (CRESS). These two groups met to discuss demands of policy makers for alternative assessment tools to be used in the nation's school systems. They noted time needed to establish validity and reliability of these new assessments, time and expenses needed to train teachers in alternative assessment procedures, and the amount of time needed to administer these new assessments (38: 275).

Additionally, in an article written by Lewis reviewing performance testing, the Educational Testing Service (ETS) related that the cost of performance tests would be higher than multiple-choice tests. The National Assessment of Educational Progress (NAEP), conducted by ETS, will include a performance task in its 1996 science assessment. The average cost of materials for this task will be six dollars per kit (23: 73).

In her study, conducted with third grade teachers in implementing classroom performance assessment projects, Shepard found that the most significant problem teachers encountered with this pilot project was time constraints. Teachers needed more time for planning, for individual assessments of their students, and for scoring individual assessments. Some teachers found

that whole-class instruction made it impossible to manage time for individual assessments of below-level readers (35: 3). Also, some teachers had difficulty differentiating between the skills to be assessed. For example, they assessed the language of a writing assignment meant to measure comprehension only (35: 4).

Current Developments in the Use of Alternative Assessments

According to a government press release on education, eight states were given research grant money in 1995 totaling twelve million dollars over four years to help develop better state assessments of student achievement. United States Secretary of Education Richard W. Riley disclosed that the federal government was making funds available under the *Goals 2000: Educate America Act* to help states put their own challenging student assessment standards in place. The eight states, chosen from twenty-seven applicants, were Delaware, Maryland, Michigan, Minnesota, North Carolina, North Dakota, Oregon, and Pennsylvania. (36: 1).

Pennsylvania was given $1,107,250 to develop and implement student assessments in science and the arts that require performance tasks, according to this press release (36: 1). Michigan and Minnesota received $584,016 and 568,142 respectively. The other six states received sums of up to $1,967,303 (36: 2).

Summary

Research of the past two decades has indicated testing policies have not had the desired effects of improving the educational system, asserted Lieberman regarding standardized tests (24: 220). In her article concerning testing policy, Darling-Hammond added that some states misuse standardized test results (15: 222).

Farr, on the other hand, expressed a desire for performance assessment to become an alternative method of student assessment. He hoped that assessment would become a means to guide students to identify their own educational strengths rather than a means to label them. Farr likewise expressed, in his "dream about assessment", that assessment would be accepted as a method to aid in teacher instruction (18: 424).

CHAPTER 3

METHODS AND PROCEDURES

Introduction

This study provided data concerning the professional convictions of teachers in Randolph County, West Virginia about standardized testing and alternatives to this testing. Because of the size of Randolph County and the profusion of different organizational patterns within the elementary schools, as well as the rural versus town environments, these data revealed varied responses and results. These varied responses and results might prove to be representative of the nation as a whole. Research questions for this study were:

1. How much time is spent in the classroom on preparation for standardized tests?

2. How much time is spent in the classroom on administration of standardized tests?

3. Do elementary teachers feel compelled to "teach to the test"?

4. How many alternative assessment methods of student evaluation do elementary teachers in Randolph County currently utilize?

5. Can these alternative assessment methods be more effective tools of student evaluation than standardized tests?

Collection of Data

After receiving permission from Mr. Larry Prichard, Superintendent of Randolph County Schools, a questionnaire was sent to each elementary school teacher in Randolph County. Refer to Appendices E, F, and G to see the request for permission, permission granted, and the teacher questionnaire. All of the eleven schools that contain elementary students were used in this data collection. They include kindergarten through five schools, kindergarten through six schools, and kindergarten through twelve schools. The kindergarten through five schools are Beverly Elementary School, Coalton Elementary School, Jennings Randolph Elementary School, Midland Elementary School, North Elementary School, and Third Ward Elementary School. George Ward Elementary School, Homestead Elementary School, and Valley Head Elementary School comprise the kindergarten through six schools. The two remaining schools are kindergarten through twelve schools. They are Harman School and Pickens School.

Three individuals who work with testing results in Randolph County were interviewed. Permission for these interviews was verbally obtained from each of the three administrators themselves. Two of the individuals interviewed are elementary school principals. One of the principals supervises a rural school, the other an urban school. The third interview was with the Director of Student Services, who is responsible for supervision of standardized testing in Randolph County. Questions for the interviews are in Appendices H and I.

Treatment of Data

After collecting teacher surveys, the data was compiled in the following manner. Yes and no question numbers one, three, four, seven, eight, ten, and eleven were graphed. Question numbers two, five, six, nine, and twelve were compared for similarities and differences. These similarities and differences were discussed.

Survey question numbers one, two, three, four, eight, and ten imparted personal opinion responses. Question five revealed how the teachers feel their students respond to standardized testing. Questions six and seven established the amount of time spent away from classroom instruction. The kinds of student assessment utilized by teachers in their respective classrooms was disclosed in survey question number nine. Question number eleven established teachers' awareness of alternative assessments and their definitions. Additional comments concerning the subject of standardized tests and alternative assessments were addressed in question number twelve. The researcher compared likenesses and differences in the answers between rural and town schools, and also, within a school. The survey also portrayed likenesses and differences between town schools and rural schools.

Talking with a principal from two different geographical locations of the county established a suitable comparison of opinions of standardized tests. The results of the director's interview was compared to the results of the principal interviews for similarities and differences.

Teacher survey answers were also compared to answers received at principal interviews. Principal interview question numbers seven and eight were compared to teacher survey question numbers three and four. Other principal interview questions compared with teacher survey questions are numbers three to one, nine to ten, and numbers ten to eleven, respectively.

Principal interview answers were compared to the Director of Student Services interview. The specific questions compared were principal question number one to director question number four. Principal question numbers five and six were compared to director question numbers six and seven. These questions were compared to ascertain likenesses and differences in tenets of the distinctive administrators.

Summary

Willie stated that there exist numerous advantages and disadvantages to both standardized tests and alternative assessments (32: 627), therefore, recommendations were made with this idea in mind. Also, the teacher survey and the principal and supervisor interviews were used to gather information on this subject of testing students.

CHAPTER 4

ANALYSIS OF COLLECTED DATA

Introduction

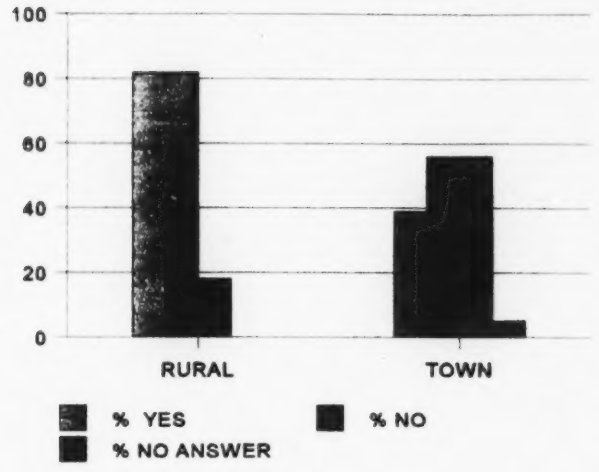
This chapter presents the data collected through elementary school teacher surveys. The graphs depict all yes and no questions, with a short explanation of percentages. Two pie charts represent advantages and disadvantages of standardized testing as stated by elementary school teachers on their surveys. The results of three interviews are also discussed and analyzed in this chapter. Interview questions and answers were compared with teacher surveys, where applicable. All the results were compared with each other in order that a conclusion may be reached and recommendations made for this study of standardized testing and alternative assessment methods.

Presentation of Data

Teacher survey question one. Were you happy with the fact that standardized testing was cancelled in all grades except third during the 1995-96 school year? *See* the graph and percentages of answers for this question on the next page.

Teacher Satisfaction or Dissatisfaction

with the Cancellation of StandardizedTests

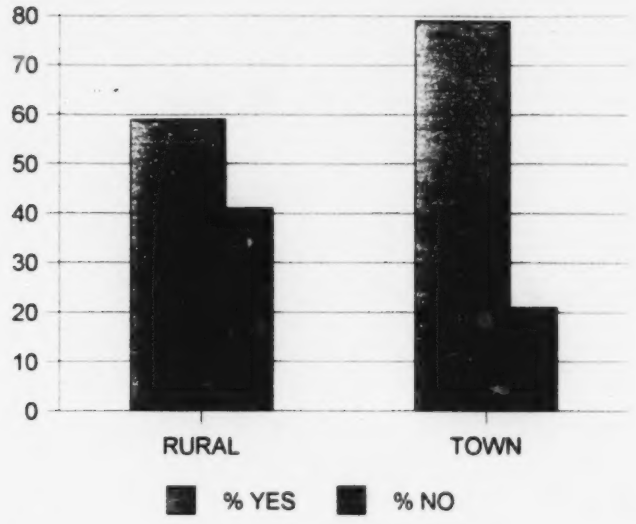


Rural schools show eighty-two percent yes, eighteen percent no. Town schools resulted in thirty-nine percent yes and fifty-six percent no. Five percent of the town school teachers were undecided. This question was one of only two that showed a difference of opinion between rural schools and town schools.

Teacher Survey Question Three. Do you feel there are advantages of standardized tests, e. g., CTBS? If you circled yes, please list one or two below.

Fifty-nine percent of rural school teachers responded yes; forty-one percent responded no. Seventy-nine percent of the town teachers responded yes; twenty-one percent responded no. See the graph on the next page which depicts these statistics derived from teacher surveys.

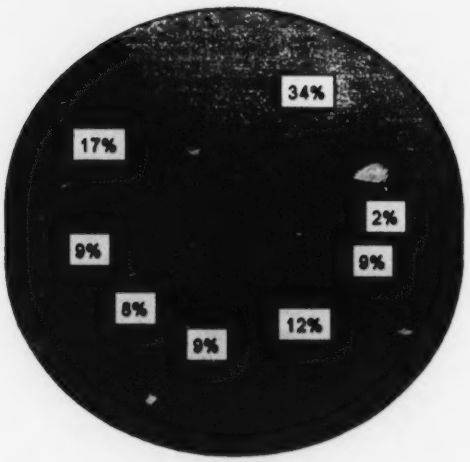
Are There Advantages to Standardized Tests?



The second part of this question asked teachers who responded yes to list one or two advantages of standardized tests. All respondents who answered yes did state at least one advantage. The researcher compiled these advantages and calculated percentages for similar items stated. These percentages are displayed on the pie chart on the next page.

Numbers on the pie chart refer to the percentages of teachers who stated similar advantages of standardized testing. The legend consists of shaded areas with one word or two word keys. These keys are defined below the pie chart. These definitions, used by the researcher, were gleaned from the teachers' surveys.

Advantages of Standardized Tests

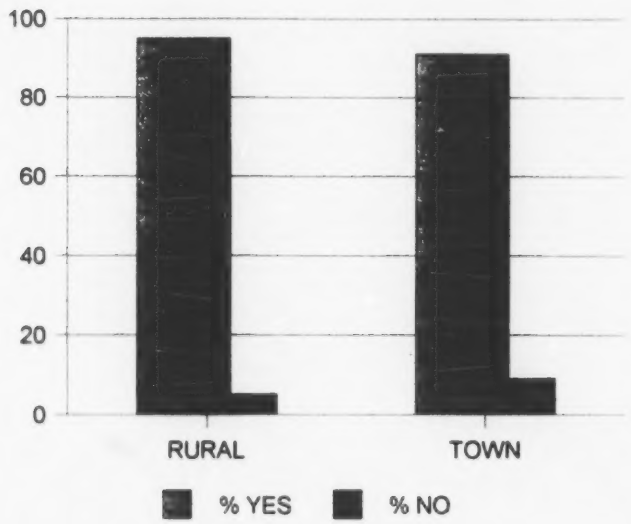


According to the teachers who answered this question, ability refers to weaknesses and strengths of students. Progress is a student's academic growth in a longitudinal study. Skill check means determining if a student's skills are on grade level. Set standards means setting teachers' standards for what to teach. Comparison refers to national norm comparison and comparison of other students of the same grade level. Curriculum means utilizing test scores for curriculum planning for the future. Test prep refers to the notion that taking standardized tests helps students prepare for other high school tests and college entrance exams. Teaching skills refers to test scores as a confirmation that teachers are teaching necessary skills for the grade level which they are currently

teaching.

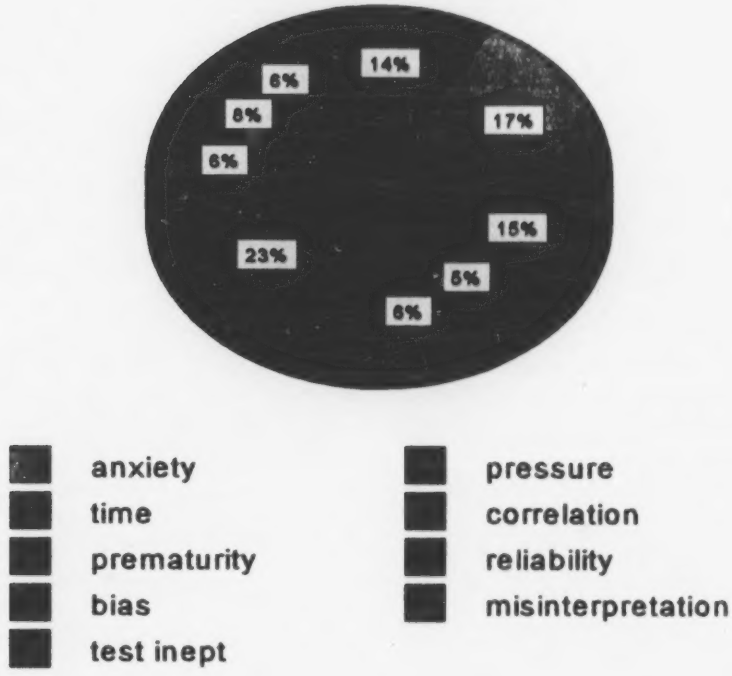
Teacher Survey Question Four. Do you feel there are disadvantages of standardized tests? If you circled yes, please list one or two below.

Are There Disadvantages to Standardized Tests?



Ninety-five percent of rural school teachers answered yes; five percent answered no. Ninety-one percent of town school teachers answered yes; nine percent answered no. The second part of this question asked teachers who responded yes to list one or two disadvantages of standardized tests. As with question three, the researcher compiled the disadvantages and calculated percentages for similar statements. The disadvantages listed by teachers on their surveys are represented on the pie chart on the next page.

Disadvantages of Standardized Tests

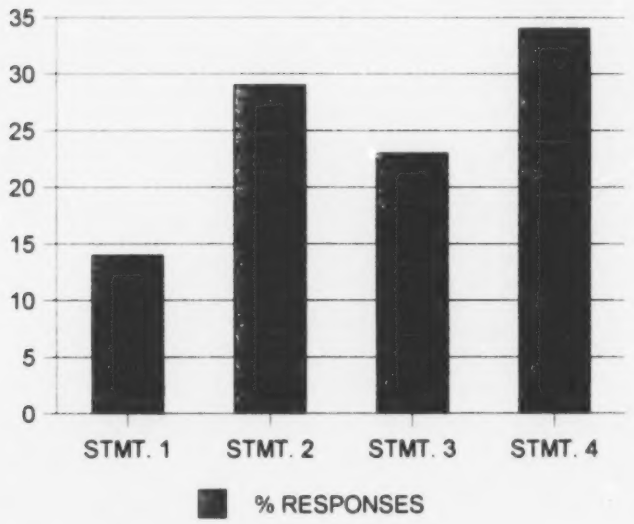


Numbers on the pie chart above refer to the percentages of teachers who presented similar statements. The legend consists of shaded areas with one word or two word keys. These keys are clarified below using the teacher definitions from their surveys.

Anxiety means student anxiety during a standardized test. Time refers to teaching time utilized for testing. Prematurity means the test is given before all skills being tested are taught in the classroom. The teachers referred to bias as cultural, socioeconomic, or academic biases concerning the format of standardized tests. Test inept means that some students do not test well. Pressure refers to the pressure teachers feel from various factions--the school

board, administration, the State Department of Education--to produce high test scores. Correlation means the tests do not correlate with the curriculum being taught. Several reliability factors were mentioned: teaching to the test, not a true assessment of student ability, and students not taking the test seriously. Misinterpretation refers to the misuse of test scores for accountability purposes.

Student Response to Standardized Tests



Teacher Survey Question Five. How do students in your classroom respond to taking standardized tests? Circle one that best fits your description or write your own description.

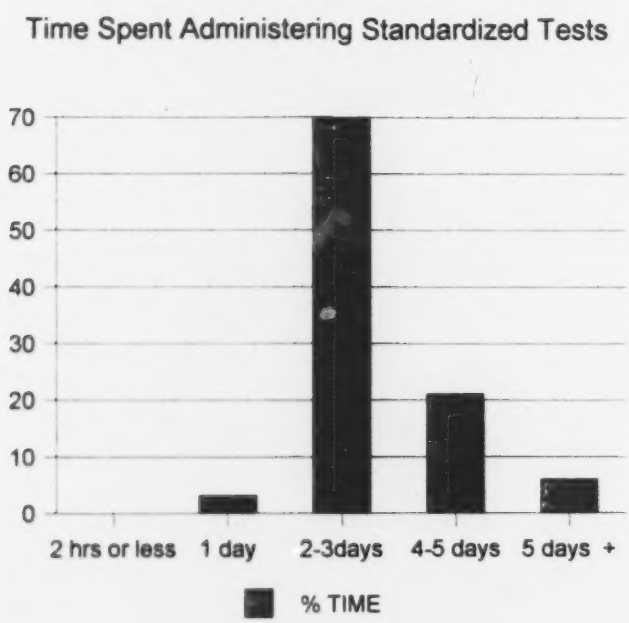
STATEMENT ONE: My students dislike taking standardized tests--fourteen percent.

STATEMENT TWO: My students don't mind taking standardized tests and try their best--twenty-nine percent.

STATEMENT THREE: Some of my students dislike standardized tests--twenty-three percent.

STATEMENT FOUR: Some of my students fill in the answer sheet without trying to comprehend the test item--thirty-four percent.

Teacher Survey Question Six. How much time do you spend administering standardized tests in your classroom?

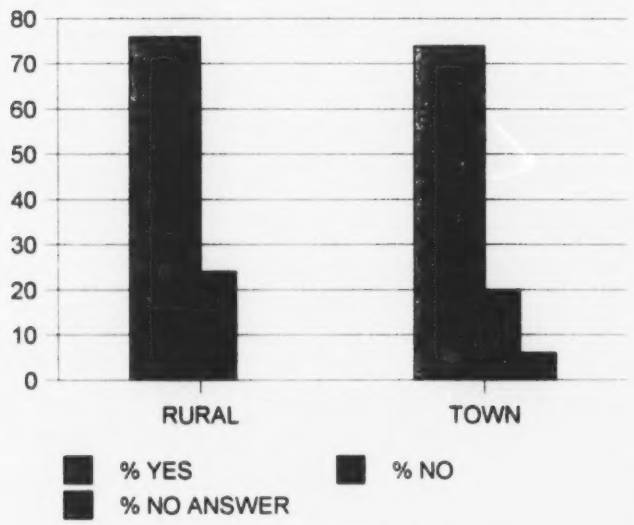


Three percent of the teachers answered that they use only one day to administer standardized tests. Seventy percent of the teachers use two to three days. Twenty-one percent test for a total of four to five days. Six percent test for five days or more.

Teacher Survey Question Seven. Do you use "practice tests" in your classroom before administering standardized tests?

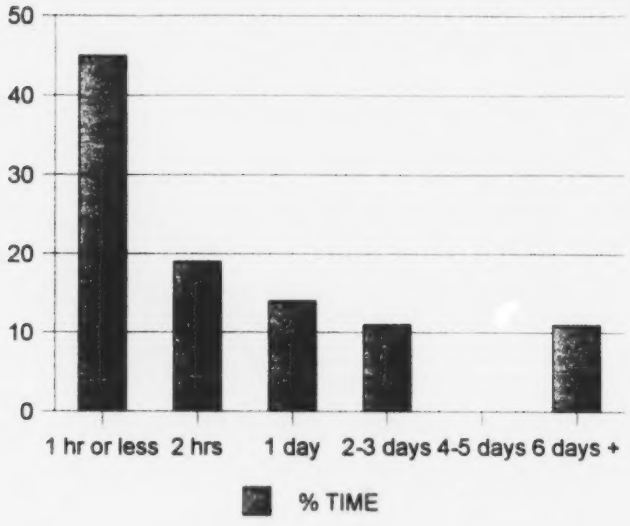
Seventy-six percent of rural school teachers answered yes to this question, and twenty-four percent of the teachers answered no. Town school surveys resulted in seventy-four percent yes, twenty percent no, and six percent of the town school teachers did not answer this question.

Do You Use "Practice Tests"?



Teacher survey question seven also asked teachers to indicate the approximate number of hours or days they spend administering "practice tests". The graph below depicts the results of the second part of question seven.

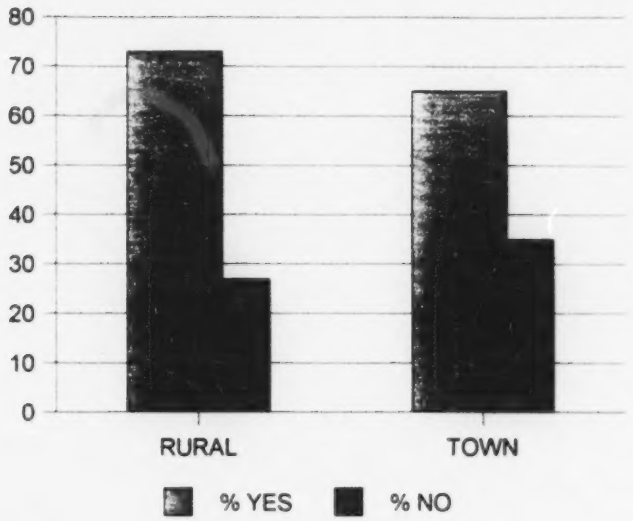
Time Spent Administering "Practice Tests"



Forty-five percent of the teachers answering question seven, stated that they used "practice tests" for standardized tests only one hour or less. Nineteen percent practice for two hours. Fourteen percent spend one day practicing. Eleven percent utilize two to three days for "practice tests", and eleven percent use six days or more for practicing. Some mention of "Test Ready" tests is made by some respondents. *See* the *Explanation and Analysis of Data* section, page forty-three, of this document for further interpretation.

Teacher Survey Question Eight. Do you feel pressure, either from yourself or from somewhere else, to teach material you know will appear on standardized tests?

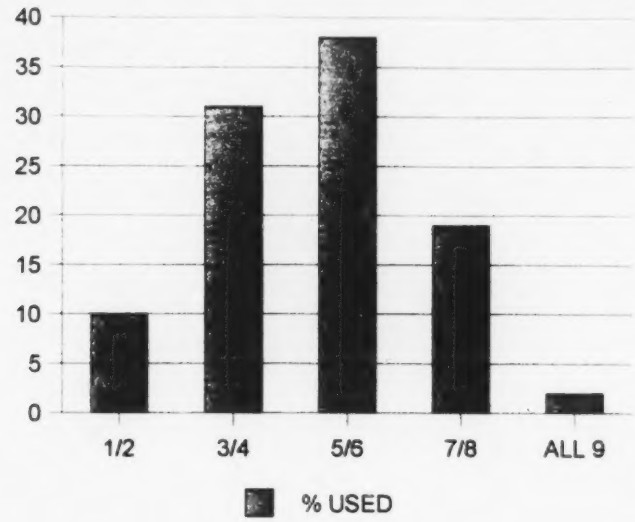
Pressure to "Teach to the Test"



Rural school teachers responded with seventy-three percent yes, twenty-seven percent no. Town school teachers responded with sixty-five percent yes, and thirty-five percent no.

Teacher Survey Question Nine. What kinds of student assessment do you use in your classroom? Circle all that pertain to your classroom and add any not listed. Spelling tests; running records; scoring rubrics; sentence dictation; anecdotal records/notes; writing vocabulary; portfolios; journals; word dictation.

Number of Student Assessments Used by Teachers



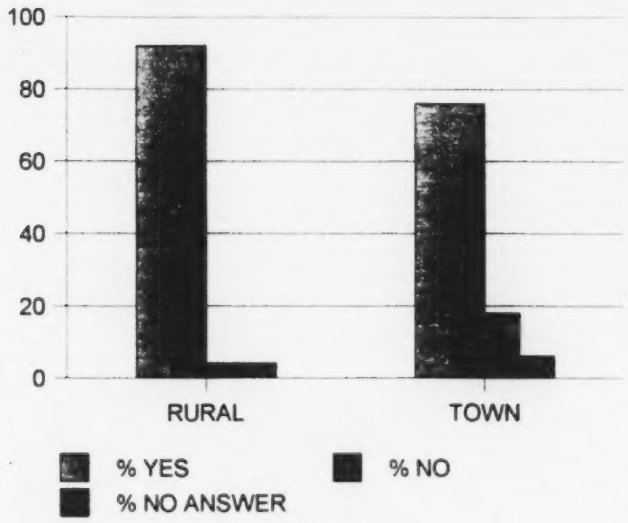
All teachers who answered the survey use at least two of the nine alternative assessments listed in question nine. The graph above displays that ten percent of the teachers use one to two of the items. Thirty-one percent use three to four of the methods. Thirty-eight percent use five to six alternative assessments. Seven to eight of these listed methods are utilized by nineteen percent, and two percent use all nine methods in their classrooms.

Twelve percent of the teachers added methods of their own not listed on the survey. Added assessment methods listed were chapter tests and quizzes, study guides, puzzles, diagrams, maps, hands-on lab activities, problem solving activities, comprehension questions, and skill check lists.

Teacher Survey Question Ten. Do you think classroom assessment methods are a better indicator of your students' progress than standardized tests?

Are Classroom Assessments a Better Indicator of

Student Progress than Standardized Tests?

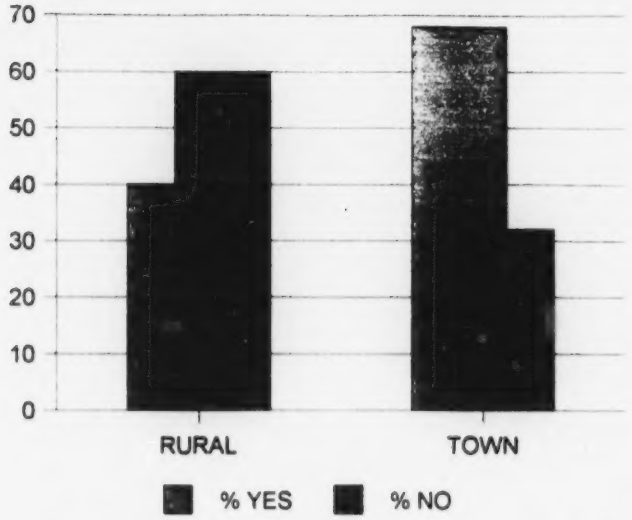


Ninety-two percent of rural school teachers answered yes to this question. Four percent answered no, and four percent left this question blank. Seventy-six percent of the town teachers answered yes. Eighteen percent answered no, and six percent left this question blank. Thirteen percent of the respondents added that both testing results and alternative assessment methods should be used in conjunction with each other.

Teacher Survey Question Eleven. Have you ever heard the term authentic or alternative assessment?

This was the second question of two that presented a difference of opinion between rural and town schools. *See* the accompanying graph on the next page.

Awareness of the Term Authentic Assessment



Forty percent of rural school teachers are familiar with the term authentic or alternative assessment. Sixty percent had not heard this term before. Sixty-eight percent of town school teachers recognized the term authentic or alternative assessment, whereas thirty-two percent did not recognize this term.

Explanation and Analysis of Data

Teacher Surveys. A total of seventy-one teacher surveys were obtained out of 108 total sent out to all elementary school teachers in Randolph County, West Virginia. Thirty-nine of these surveys were returned from town schools. Thirty-two surveys were returned from teachers of rural schools. Only two of these teacher survey questions, question one and question eleven, presented a difference of opinion between town schools and rural schools.

A total majority of yes answers surfaced for teacher question number one;

even though a majority of rural schools answered yes, and a majority of town schools answered no. This was only one of two questions on the teacher survey that posed a difference of opinion between rural schools and town schools. The results of this question are depicted as percentages in the graph shown on page thirty in the *Presentation of Data* section of this chapter. This majority of yes answers was consistent with the principal answers to principal question number three. Both of these teacher and principal questions concerned their feelings for standardized testing of elementary school students. Of all schools polled, only one town school was inconsistent with the majority. This town school answered nine to one for being unhappy that standardized tests were cancelled during the 1995-96 school year. One respondent answered he or she had mixed feelings about the subject. This town school consists of a majority of high socioeconomic families, which may account for their answers to this question. Both principals and a majority of elementary classroom teachers agreed that standardized tests alone do not measure a student's abilities. Teacher question number ten revealed that the majority of teachers were in agreement that classroom evaluative methods are a better indicator of students' abilities. *See* the graph pertaining to this question on page forty-one in the *Presentation of Data*. Again, the only exception was the town school with a majority population of high socioeconomic families. This town school's answer to question number ten was split almost evenly, with six no answers and five yes answers. This school's no answers were qualified with a statement noting that they thought standardized

test results were comparable in importance with classroom teacher evaluative methods.

When asked about their knowledge of the term "authentic assessment", neither principal was aware of the term, but understood the methods because their teachers use these methods. The majority of teachers in town schools were aware of the term "authentic assessment"; the majority of teachers in rural schools had not heard of this term. Whether the teachers knew the term or not, all teachers answered that they use at least one of the authentic assessment measures listed on the survey in their classrooms. As represented in the graph on page forty-two, question number eleven was only one of two questions that showed a difference of opinion between town schools and rural schools.

Teacher survey question number seven pertained to the amount of classroom time utilized in practicing for standardized tests. The majority of teachers used only the practice sections of the standardized tests for actual practicing. The time involved ranged from fifteen minutes to two hours. A minority used "Test Ready" tests to practice. One teacher indicated that the principal expected the teachers to use this practice test in their school. The time involved in using "Test Ready" was four to five days. A total of four teachers practiced for five days or more. One of these teachers said he or she practiced a total of one to two days per month. Another teacher practiced for two weeks. Percentages for each of the time ranges may be obtained from the graph found on page thirty-eight in the *Presentation of Data* segment of this article.

Question number six of the teacher survey regarded amounts of time

teachers spend administering the standardized test. These answers were consistent from one school to another. The majority of teachers spend two to three days total classroom time, using only mornings to administer the test. One teacher indicated using only one day for administering the test. Two teachers indicated using five days or more. *See* the graph for percentages of the time ranges for question six on page thirty-six.

Teachers were asked to relate student responses to standardized tests in question number five of the survey. Two statements out of four yielded greater responses, totaling thirty-nine out of sixty-one. These statements are "My students don't mind taking standardized tests and try their best" and "Some of my students fill in the answer sheet without trying to comprehend the test item." One teacher enumerated these statements by commenting that "successful students try; low achievers get frustrated and give up." The percentages of all four of the statements used in this survey question can be found on the graph in the *Presentation of Data* on page thirty-five.

Standardized tests were not administered in all elementary grades during the 1995-96 school year. Question number two asked the teachers for reasons why they thought this occurred. Every teacher except five answered with financial reasons. Two teachers thought the test was cancelled because of the number of snow days Randolph County had during the 1995-96 school year. One teacher supposed there were too many complaints about the test the year before, and two teachers assumed the reason was a change in teaching trends.

Principal Interviews. The same questions were asked of both a rural

school principal and a town school principal. The majority of their answers were similar, with the exception of numbers five, seven, and eight. These three questions are analyzed below.

Question number five asked about the use of "practice tests" before administering standardized tests in their schools. The rural principal replied his teachers will use "practice tests" for the first time in the 1996-97 school year. This must be done, he stated, because his school's standardized test score results did not meet state mandates the previous school year. He further stated that using "practice tests" cost his school $200.00 to $300.00, which could have been better utilized on classroom materials (43: np). The town principal's response to this question regarding "practice tests" was that the teachers feel more comfortable with administering standardized tests if the children have had practice taking a similarly constructed test (46: np).

Question number seven asked the principals to state one major advantage of standardized tests. The rural principal considered a major advantage of standardized tests to be the presentation of a student's strengths and weaknesses in a longitudinal study (43: np). This principal's answer was consistent with a majority of teacher answers to teacher survey question number three, in which teachers were asked what they considered to be an advantage of standardized tests. Whereas, the town principal indicated a need for national comparison to be a significant advantage for testing (46: np). This answer also appeared in answers to teacher survey question number three, but was not as predominant as the answer referring to students' strengths and weaknesses.

See pie chart, page thirty-two: labels "ability" and "comparison".

Both principals were asked to relate one major disadvantage of standardized tests in question number eight. The rural principal expressed concern that standardized testing used valuable instructional classroom time (43: np). A few teachers also mentioned this as a disadvantage to standardized tests in teacher survey question number four. The age of the standardized test being used, and the times of the standardized test in the school year were considered to be disadvantages to the town principal. This principal explained that standardized tests are administered before all the skills that appear on the test are taught (46: np). Approximately the same number of teachers responded similarly with the town principal's response, as they did with the rural principal's response. *See* the pie chart on page thirty-four: labels "time" and "anxiety".

The remaining eight questions asked of the principals revealed very similar answers. They both utilize standardized test scores to improve deficient areas of their curriculum. Neither principal utilizes test scores alone to make a life-changing decision for any student. Both principals recommend using teacher - assessments along with standardized test results to evaluate students.

Both the rural and the town principal agreed that standardized testing is a stressful situation, and therefore not a valid tool. The rural principal regarded standardized tests as a necessary tool for assessment, but stated that the results are being misused for accountability of teachers and schools instead (43: np). This reason was also brought to notice by teachers in survey question number four, which was the reason the rural principal gave for the teachers' dislike of

standardized tests. The town principal stated that the teachers dislike the standardized test format, because it does not conform to their teaching techniques (46: np). Many teachers mentioned this rationale as a disadvantage of standardized tests. *See* the pie chart, page thirty-four: labels "misinterpretation" and "correlation".

Both principals admitted that they encourage their teachers to teach to the test. Their reason for doing this was accountability. Teacher survey question number eight confirms this, as the majority of teachers answered that they feel pressure to "teach to the test". *See* the pie chart, page thirty-four, labeled "reliability".

Neither rural nor town principal was aware of the term "authentic assessment", but teachers utilize various types of authentic assessment in both their schools. It is of interest to note here, that rural school teachers were unaware of the new term; town school teachers did recognize the term. However, teachers from all schools, utilized at least three to four alternative assessment methods in their classrooms, whether or not they knew the terminology. *See* the graph on page forty.

Director of Student Services Interview. This interview revealed some similar, as well as dissimilar, answers as compared to the principal interviews. The first similarity is found in question number four. The county office director agreed with the principals in that standardized tests provide a measure of curriculum and instructional goals. He explained further that standardized tests

are mandated by the State Department of Education, and that there is a task force who writes test items to correspond with the Education Department's instructional goals. However, he disagreed with the principals in his answer to this same question concerning accountability issues dealing with teacher performance (47: np). Both principals agreed that accountability should not be an objective of standardized tests; whereas, the county office director proffered accountability as one of the test's objectives.

Question numbers six and seven deals with "teaching to the test" and "practice tests". The director confirmed that he was aware of teachers "teaching to the test". He is also aware that some schools use "practice tests" to improve standardized test scores. He agreed that this use of practice tests and teaching to the test "skews" the results; however, it is his belief that the positive aspect of measuring instructional goals outweighs this detriment (47: np).

Another similarity between the director's and the principals' answers pertains to standardized test results. The director concurred, in question number eight, that the results should not be used alone as a measure of a student's ability. The results should be used in conjunction with other teacher assessments (47:np).

When asked if there are other types of standardized tests available in addition to multiple-choice, the director replied that there are alternatives to choose from. He explained that there are two types. One type emphasizes basic skills. The other type focuses on higher order thinking skills. He reported that Randolph County would be administering a new standardized test this year--the

Stanford Achievement Test, rather than the Comprehensive Test of Basic Skills. He revealed that the Stanford Achievement Test is a standardized test of higher order thinking skills (47: np).

Summary

The data in this study suggests a need by most of the participants for standardized tests of elementary school students. Even those teachers who dislike administering standardized tests feel a need for the use of its results. Many participants, however, have a concern about the misuse of test results. Many respondents to the teacher survey and one principal conveyed concern that test results were misused for accountability, rather than used as a helpful evaluative tool for their students.

CHAPTER FIVE

SUMMARY AND CONCLUSIONS

Introduction

This chapter presents recommendations for the use of standardized test results. The recommendations have been made after careful examination of the teacher surveys returned and the results of the administrative interviews. In examination of this information, some answers to the problem presented in Chapter One of this document have been explored.

Conclusions

Restatement of the Problem. Should standardized tests be replaced with or used in conjunction with alternative assessments in order to be a more effective measurement tool for evaluating students' abilities in elementary schools? The majority of Randolph County elementary school teachers believe that standardized tests should not be replaced, but used in conjunction with other teacher evaluative methods. The teacher survey results recorded on the graph, page thirty-one and the pie chart, page thirty-two, verify this. These results have occurred despite the fact that the majority of teachers dislike standardized tests, as indicated on the graph on page thirty. These results could be comparable to the nation for two reasons. First, Randolph County presents a wide diversity of both geographics and economics in relation to location and types of schools and status of school population. Second, a fair representation of information for this study was received from each of these diverse areas.

Restatement of the Research Questions and their Answers.

1. How much time is spent in the classroom on preparation for standardized tests? The majority of Randolph County elementary students spend one hour or less practicing for standardized tests.

2. How much time is spent in the classroom on administration of standardized tests? The majority of Randolph County elementary teachers spend two to three school days administering standardized tests in their classrooms.

3. Do elementary teachers feel compelled to "teach to the test?" The majority of Randolph County elementary teachers answered yes.

4. How many alternative assessment methods of student evaluation do elementary teachers in Randolph County currently use? The majority of Randolph County teachers use five to six alternative assessment methods in their classrooms.

5. Can these alternative assessment methods be more effective tools of student evaluation than standardized tests? The majority of Randolph County elementary teachers answered yes.

Recommendations

According to the results of this study, three recommendations should be made regarding standardized tests. First, said tests should be utilized in conjunction with other student evaluation methods chosen by the classroom teacher. Second, standardized tests should be revised to include more real-life tasks and more authentic types of assessments. Third, standardized tests should

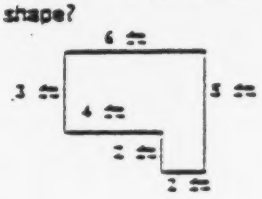
be written to correlate with the curriculum it is measuring.

Summary

Although the practice is not always acceptable by all those involved, testing of elementary students is necessary. This conclusion has been reached after careful examination of the teacher survey results, along with the administrator answers, procured from interviews. Even teachers who object to administering standardized tests in their classroom, listed advantages of standardized tests when answering their survey. Although the principals interviewed had different feelings on some of the issues of standardized testing, they both felt it necessary to utilize test scores in some manner. The Director of Student Services supported this conclusion as well. Because of the diversity of this county, results of this study could be analogous to the nation as a whole concerning this subject of standardized testing and alternative assessment methods.

Appendix A

This is an example of questions on a third grade standardized achievement test in mathematics (30: 27).

Computation	Concepts and Applications
Add. 1. 3842 + 4104 ○ 7946 ○ 7746 ○ 7906 ○ 7942 ○ None of these	1. What is 763 rounded to the nearest hundred? ○ 700 ○ 750 ○ 760 ○ 800
Subtract. 2. 82 - 3 = ○ 89 ○ 79 ○ 81 ○ 52 ○ None of these	2. How much change will you get if you have $6.55 and spend $4.32? ○ $2.23 ○ $2.43 ○ $3.23 ○ $10.87
Multiply. 3. 6 x 9 ○ 63 ○ 48 ○ 54 ○ 69 ○ None of these	3. What is the perimeter of this shape?  ○ 20 ft ○ 21 ft ○ 22 ft ○ 23 ft
Divide. 4. 8 ⟌ 16 ○ 2 ○ 24 ○ 20 ○ 8 ○ None of these	

BEST COPY AVAILABLE

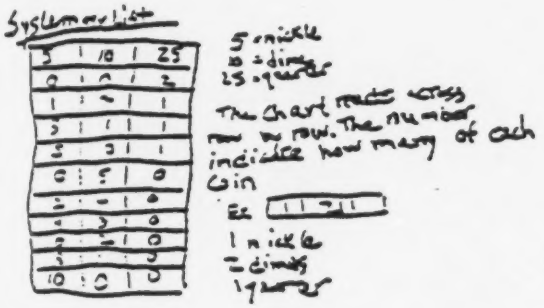
Appendix B

This is an example of third grade performance assessment questions in mathematics (30: 29).

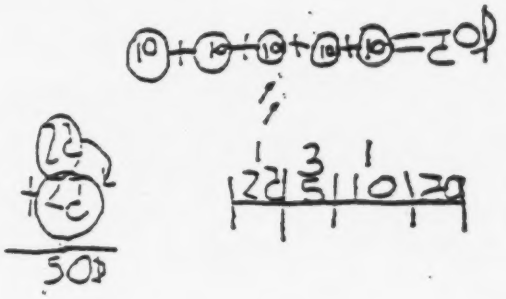
1. Bus Ride – A friend of yours, who just moved to the United States, must ride the bus to and from school each day. The bus ride costs 50 cents. Your friend must have exact change and must use only nickels, dimes, and quarters. Your friend has a problem because she does not yet understand our money, and she does not know how to count our money.

Help your friend find the right coins to give to the bus driver. Draw and write something on a whole sheet of paper that can help her. She needs a sheet of paper that can show which combinations of coins can be used to pay for the 50-cent bus ride.

Sample Student Answer 1

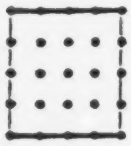


Sample Student Answer 2



BEST COPY AVAILABLE

BEST COPY AVAILABLE

2. 

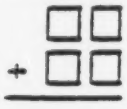
For the figure at left, show 1/2 in as many ways as you can. You may draw more figures, if necessary. For each way you find, explain how you know you have 1/2.

3. Suppose you couldn't remember what 8 x 7 is. How could you figure it out?

4. Our class of 25 students is going to the Denver Art Museum. How many cars do we need if 4 students can go in each car? How many do we need if only 3 students can go in each car?

5. Adam says that 4 - 52 is 452. Is he right or wrong? What would you tell Adam?

6. Put 4 different one-digit numbers in the boxes to make the largest possible answer.



How did you know what to choose?

BEST COPY AVAILABLE

Appendix C

This is an example of an alternative assessment used for a science experiment (4: 44).

What Happens When You Put Water on These Things

ACTIVITY TO CONDUCT

1. Place a drop of water on each material.

RECORD FINDINGS

2. Look carefully. What do you see? Write down what happens to the water on each of the materials.
   - A. Plastic: nothing happens
   - B. Painted wood: nothing happens
   - C. Brick: It fades so you can't see it
   - D. Metal: the drop becomes a circle
   - E. Roof shingle: It fades so you can't see it
   - F. Glass: It stays the same
3. Now use your magnifying glass and look at each material very closely.
4. Look at the material in the plastic bag very closely. Do not open the bag.
5. Write down what you think would happen if you put a drop of water on the material.

FORMULATE HYPOTHESIS

It would soak throuh

ACCOUNT FOR HYPOTHESIS

6. Write down why you think this will happen.

Because it soaked throuh the Brick, and Roof shingle, and it is made of the same specimens.

BEST COPY AVAILABLE

Appendix D

This is an example of an alternative assessment used for writing (4: 66).

Parent Portfolio Review[13]

Reader's Name _____ Date _____

Student's Name _____

Please read everything in your child's writing folder, including drafts and commentary. Each piece is set up in back-to-front order, from rough draft to final copy. Further, each piece is accompanied by both student and teacher comments on the piece and the writing process. Finally, the folders also include written questionnaires where students write about their strengths and weaknesses as writers.

We believe that the best assessment of student writing begins with the students themselves, but must be broadened to include the widest possible audience. We encourage you to become part of the audience.

When you have read the folders, please talk to your children about their writing. In addition, please take a few minutes to respond to these questions.

- Which piece of writing in the folder tells you the most about your child's writing?
- What does it tell you?
- What do you see as the strengths in your child's writing?
- What do you see as needing to be addressed in your child's growth and development as a writer?
- What suggestions do you have that might aid the class's growth as writers?
- Other comments, suggestions?

Thank you so much for investing this time in your child's writing.

BEST COPY AVAILABLE

Appendix E

Rt. 3 Box 433
Elkins, West Virginia
August 12, 1996

Mr. Larry Prichard
40 Eleventh Street
Elkins, West Virginia

Dear Mr. Prichard:

I am writing to request your permission to send the elementary teachers of Randolph County a survey. I am currently writing a thesis to complete my Master's Degree for Salem-Teikyo University. I need the results of the teacher survey to complete my thesis. The teacher survey concerns student assessment methods used in the classroom and standardized testing.

Thank you for your timely consideration of this matter.

Sincerely,

Judith L. Sinkule

Appendix F

Survey of Student Evaluation Methods used by Elementary School Teachers

1. Were you happy with the fact that standardized testing was canceled in all grades except third during the 95-96 school year?

Yes No

2. What do you feel was the reason standardized tests were not given in 95-96?

3. Do you feel there are advantages of standardized tests, e.g., CTBS?

Yes No

If you circled yes, please list one or two below.

4. Do you feel there are disadvantages of standardized tests?

Yes No

If you circled yes, please list one or two below.

5. How do students in your classroom respond to taking standardized tests? Circle one that best fits your description or write your own description.

My students dislike taking standardized tests.

My students don't mind taking standardized tests and try their best.

Some of my students dislike standardized tests.

Some of my students fill in the answer sheet without trying to comprehend the test item.

6. How much time do you spend administering standardized tests in your classroom? (List hours or number of days or half days, etc.)

7. Do you use "Practice tests" in your classroom before administering standardized tests?

Yes No

If you circled yes, please indicate the approximate number of hours or days (or half days) you spend administering "practice tests".

8. Do you feel pressure, either from yourself or from somewhere else, to teach material you know will appear on standardized tests?

Yes No

9. What kinds of student assessment do you use in your classroom? Circle all that pertain to your classroom and add any not listed below.

Spelling tests	Sentence dictation	Portfolios
Running records	Anecdotal records/notes	Journals
Scoring rubrics	Writing vocabulary	Word dictation

10. Do you think the methods above are a better indicator of your students' progress than standardized test results?

Yes No

11. Authentic or alternative assessment is a type of student assessment that measures "real life" tasks and situations rather than just basic facts of information, as standardized tests do. (Some forms of this type assessment are listed in #9 above.) Have you ever heard the name authentic or alternative assessment before?

Yes No

12. If you have any other comments you'd like to add concerning this topic of student evaluation methods, please feel free to do so.

Appendix G

Randolph County Schools

40 ELEVENTH STREET, ELKINS, WV 26241
Telephone (304) 636-9150
FAX (304) 636-9157

LARRY G. PRICHARD
Superintendent of Schools

Edward Tyre, President
Mark Rizzie, Vice President
William HC Rice
Mary Alice Bester
Sherwood Collett

August 19, 1996

Mrs. Judith L. Sinkule
Route 3, Box 433
Elkins, WV 26241

Dear Mrs. Sinkule:

I have reviewed your request to send out a survey to elementary teachers in Randolph County and I concur with the request. It is a good instrument. Would you please share the results with me and Ms. Barbara Korn upon completion of this survey.

Thank you for your interest in our school system and I congratulate you for your pursuit of a Masters Degree in Education.

Should I be able to assist you in any way, please don't hesitate to call.

Sincerely,

Larry G. Prichard, Superintendent
RANDOLPH COUNTY SCHOOLS

LGP/cah

cc:

BEST COPY AVAILABLE

Education Today For A Better Tomorrow

BEST COPY AVAILABLE

Appendix H

Principal Interview of Student Evaluation Methods used in their Schools

1. How do you utilize the results of standardized tests, e. g., CTBS?

Do you ever use these results for a life-changing decision regarding a student, for example retention or promotion?

2. What kinds of student assessment do your teachers use throughout the school year?

3. Do you like using standardized tests in your school? Why or why not?

4. Do your teachers like using standardized tests in their classrooms? Why or why not?

5. Do you use "Practice Tests" before giving standardized tests? Why or why not?

6. Do you encourage your teachers to teach material that will appear on standardized tests?

7. What do you feel is one major advantage of standardized tests?

8. What is one major disadvantage of standardized tests?

9. Do you think standardized test results used alone is a valid indicator of a student's ability? If not, what else do you recommend for student assessment?

10. Authentic or alternative assessment is a type of student assessment that measures "real life" tasks and situations rather than just basic facts of information, as standardized tests do. Have you ever heard this term before?

11. Some examples of authentic assessment are portfolios, journals,

projects, anecdotal notes or records, running records. Are you aware of many of your teachers using these types of assessments?

Appendix I

Interview for Director of Student Services

1. Why must we give standardized tests in the elementary grades?
2. Who mandates standardized tests? Why?
3. Who determines what type of standardized test to be used?
4. How do you utilize standardized test results?
5. Why are multiple-choice standardized tests used? Are there other types available?
6. Are you aware of any schools "teaching to the test" or using "practice tests"?
7. Do you think using "practice tests" skews standardized test scores?
8. Do you think standardized tests alone accurately measure a student's true ability?
9. Which standardized test is Randolph County using this year?
10. Why did the county change from Comprehensive Test of Basic Skills to the Stanford-Binet Test?

BIBLIOGRAPHY

Books

1. Ebel, Robert L. The Uses of Standardized Testing. Bloomington, Phi Delta Kappa Educational Foundation, 1977.
2. Educational Testing Service Trustees' Colloquy. Different Needs Difficult Answers. Princeton: ETS, 1995.
3. Gould, Stephen Jay. The Mismeasure of Man. New York: Norton, 1981.
4. Hart, Diane. Authentic Assessment: A Handbook for Educators. Menlo Park: Addison-Wesley, 1994.
5. Perrone, Vito. The Abuses of Standardized Tests. Bloomington, Phi Delta Kappa Educational Foundation, 1977.
6. Rothman, Robert. Measuring Up: Standards, Assessment, and School Reform. San Francisco, Jossey-Bass, 1995.

Periodicals

7. Afflerbach, Peter, ed. "Teachers' Choices in Classroom Assessment." Reading Teacher 48 (Apr. 1995): 622-24.
8. Bracey, Gerald W. "The Impact of Testing on Teaching (Concluded?)." Phi Kappan 75 (June 1994): 806.
9. Brainard, Gladys J. "This is a Test." The Delta Kappa Gamma Bulletin 56-2 (Winter 1990): 59.
10. Brandt, Ron. "On Restructuring Schools: A Conversation with Fred Newmann." Educational Leadership 53 (Nov. 1995): 70-3.
11. Campbell, N. Jo. "Interpreting Scores from Standardized Tests." The

Clearing House 67 (Jul/Aug 1994): 314-16.

12. Carver, Ronald P. "What do Standardized Tests of Reading Comprehension Measure in Terms of Efficiency, Accuracy, and Rate?" Reading Reasearch Quarterly 27, no. 4 (1992): 347-59.
13. Cizek, Gregory J. "The Big Picture in Assessment and Who Ought to Have It." Phi Delta Kappan 77 (Nov. 1995): 246-49.
14. Clinchy, Evans. "America 2000, Reform, Revolution, or Just More Smoke and Mirrors?" Phi Delta Kappan 73, no. 3 (Nov. 1991): 185-91.
15. Darling-Hammond, Linda. "The Implications of Testing Policy for Quality and Equality." Phi Delta Kappan 73, no.3 (Nov. 1991): 220-25.
16. Denman, Carrie. "Writers, Editors, and Readers: Authentic Assessment in the Newspaper Class." English Journal 84 (Nov. 1995): 55-57.
17. Doyle, Denis P. "America 2000." Phi Delta Kappan 73, no.3 (Nov. 1991): 185-91.
18. Farr, Roger. "I Have a Dream about Assessment." Reading Teacher 49 (Feb. 1996): 424.
19. Hynes, William A. "A Citizen's Remonstrance." Phi Delta Kappan 75 (Mar. 1994): 559-62.
20. Kaufhold, John A. "Testing, Testing." The American School Board Journal 182 (Jan. 1995): 41-42.
21. Khattri, Nidhi, Michael B. Kane, and Alison L. Reeve. "How Performance Assessments Affect Teaching and Learning." Educational Leadership 53 (Nov. 1995): 80-83.

22. Kim, JinGyu, James E. McLean, and Asghar Iran-Nejad. "The Integration of Teaching, Testing, and Learning." Educational Research Quarterly 19 (Mar. 1996): 29-50.

23. Lewis, Anne C. "Performance Testing." Education Digest 61 (Dec. 1995): 73.

24. Lieberman, Ann. "Accountability as a Reform Strategy." Phi Delta Kappan 73, no. 3 (Nov. 1991): 219-20.

25. Madaus, George F. "The Effects of Important Tests on Students." Phi Delta Kappan 73, no. 3 (Nov. 1991): 226-31.

26. Madaus, George F. "Test Scores as Administrative Mechanisms in Educational Policy." Phi Delta Kappan 66 (May 1985): 611-17.

27. Salganik, Laura Hersh. "Why Testing Reforms are so Popular and How They are Changing Education." Phi Delta Kappan 66 (May 1985): 607-10.

28. Savage, David. "Scrutinize Students' Test Scores, and They Might not Look so Rosy." The American School Board Journal 171 (Aug. 1984): 21-24.

29. Sewall, Gilbert T. "America 2000, an Appraisal." Phi Delta Kappan 73, no.3 (Nov. 1991): 204-09.

30. Shepard, Lorrie, and Carribeth L. Bliem. "Parents' Thinking about Standardized Tests and Performance Assessments." Educational Researcher 24 (Nov. 1995): 25-32.

31. Webb, C. Anne. "No to National Standards." English Journal 84 (Dec. 1995: 87).

32. Willie, Charles V. "The Problem of Standardized Testing in a Free and

Pluralistic Society." Phi Delta Kappan 66 (May 1985): 626-28.

Public Documents

33. U.S. House of Representatives. Hearing on the Office of Educational Research and Improvement. Hearing before the Subcommittee on Select Education of the Committee on Education and Labor. 102nd Cong., 1st sess., 23 Apr. 1991. Washington: GPO, 1991.

Pamphlets

34. Bagin, Carolyn B. and Lawrence M. Rudner. "What Parents Know about Standardized Testing in Schools." ERIC Clearinghouse on Assessment and Evaluation. Sept. 21, 1993.

Computer Materials

35. Shepard, Lorrie A. Using assessment to Improve Learning. Microsoft Windows, Netscape, Infoseek. July 8, 1996. (http://www.enc.org/online/ENC2404/2404.html).

36. Thomas, David, "contact". States to Explore Better Ways of Student Testing. Microsoft Windows, Netscape, Infoseek. Nov. 28, 1995. (http://www.ed.gov/PressReleases/11-1995/assess.html).

ERIC Documents

37. James, Julia Carol and C. Kenneth Tanner. "Standardized Testing of Young Children." Journal of Research and Development in Education 26, no. 3 (Spring 1993): 143-50. ERIC. EJ471021.

38. Maeroff, Gene I. "Assessing Alternative Assessments." Phi Delta Kappan

73, no. 4 (Dec. 1991): 272-81. ERIC. EJ435701.

39. Morehead, Michael A. and Lawrence Lyman. "Teacher Selection: The Misuse of Tests and Other Criteria Abound." Contemporary Education 59, no. 3 (Spring 1990): 159-61. ERIC. EJ377030.

40. Paris, Scott G. et. al. "A Developmental Perspective on Standardized Achievement Testing." Educational Researcher 20, no. 5 (Jun.-Jul. 1991): 12-19. ERIC. EJ432070.

41. Popham, W. James et. al. "Measurement-Driven Instruction: It's on the Road." Phi Delta Kappan 66, no. 6 (May 1985): 628-34. ERIC. EJ 323010.

42. Urzillo, Robert L. "Competency Testing: Blessing or Bane?" Contemporary Education 59, no. 1 (Fall 1987): 13-14. ERIC. EJ360040.

Unpublished Sources

43. Ferdinand Baldt. Interview by author. 12 Aug. 1996.

44. Kim, JinGyu. "Three Approaches for the Integration of Teaching, Testing, and Learning." Paper presented at the Meeting of the Mid South Educational Research Association, 21st Annual Conference. Knoxville, Tenn., 11 13 Nov. 1992. ERIC. ED354250.

45. Randolph County Update of Comprehensive Educational Facilities Plan. Clifford Wilmoth, Director. Submitted Nov. 1, 1993. Randolph County Board of Education, Elkins.

46. Rebecca Whiteman. Interview by author. 20 Aug. 1996.

47. Victor Thompson. Interview by author. 12 Aug. 1996.